



Europäisches Patentamt
European Patent Office
Office européen des brevets



⑪ Publication number : **0 438 194 A2**

⑫

EUROPEAN PATENT APPLICATION

⑳ Application number : **91200042.9**

⑤① Int. Cl.⁵ : **G06F 15/72**

㉔ Date of filing : **11.01.91**

③① Priority : **15.01.90 GB 9000842**

④③ Date of publication of application :
24.07.91 Bulletin 91/30

⑧④ Designated Contracting States :
DE FR GB IT

⑦① Applicant : **PHILIPS ELECTRONIC AND
ASSOCIATED INDUSTRIES LIMITED**
Philips House 188 Tottenham Court Road
London W1P 9LE (GB)
GB
Applicant : **N.V. Philips' Gloeilampenfabrieken**
Groenewoudseweg 1
NL-5621 BA Eindhoven (NL)
DE FR IT

⑦② Inventor : **Winser, Paul Anthony**
c/o Philips Research Laboratories
Redhill, Surrey RH1 5HA (GB)

⑦④ Representative : **Andrews, Arthur Stanley et al**
PHILIPS ELECTRONICS Patents and Trade
Marks Department Philips House 188
Tottenham Court Road
London W1P 9LE (GB)

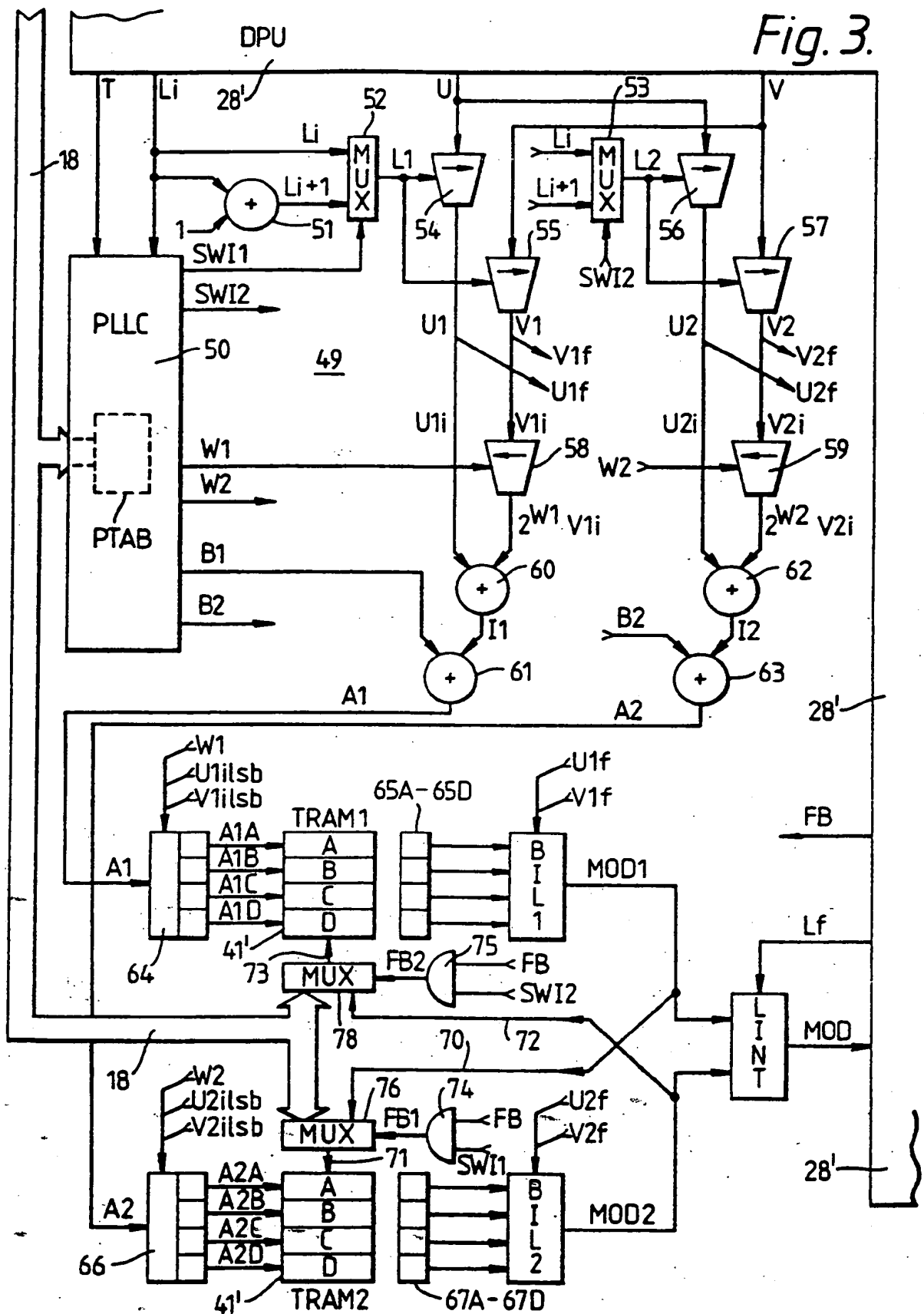
⑤④ Display apparatus.

⑤⑦ A display apparatus for real-time 3-D image synthesis comprises a display processor (28',49') having associated texture memory (41') for storing pyramidal arrays of texture element ("texel") values. Each pyramid array represents a 2-D modulation pattern at several distinct levels of resolution. The apparatus further includes an inter-level interpolator (LINT) for combining data (MOD1 and MOD2) from two levels of the pyramidal array. The texture memory (41') is divided into two banks (TRAM1 and TRAM2) and the apparatus is arranged to store successive levels of a pyramidal array alternately in the first and second banks (TRAM1 and TRAM2) of the texture memory (41') so as to allow parallel readout (at 65A-65D and 67A-67D) of data from any two adjacent levels of the pyramidal array for supply to the inter-level interpolator (LINT).

EP 0 438 194 A2

Jouve, 18, rue Saint-Denis, 75001 PARIS

Fig. 3.



DISPLAY APPARATUS

The invention relates to a display apparatus comprising a display processor having associated display memory and texture memory for storing at least one pyramidal or part-pyramidal array of texture element ("texel") values representing a two-dimensional (2-D) modulation pattern at at least two distinct levels of resolution identifiable by respective values of a level coordinate, the apparatus further comprising inter-level interpolating means responsive to a fractional part of a received level coordinate for combining together corresponding texel values from at least two levels of a stored pyramidal array so as to generate an interpolated texel value.

Such an apparatus is described in WO 85/00913 and provides real-time synthesis and display of images representing three-dimensional scenes and objects for flight simulation.

The apparatus implements a technique known in the art as "texture mapping", in which a 2-D pattern (the "texture") is pre-generated and stored in the texture memory, whereupon a single primitive is then rendered (transformed from "object space" into "screen space" and scanned into the display memory) with the texture mapped onto it. The technique enables a large amount of surface detail to be represented without a corresponding increase in the number of primitives that have to be rendered to produce an image. In a simple case, the stored pattern defines the colour of an object's surface, so that the texel values may constitute the colour values which may be written directly into the display memory. In a more general case, the texel values may be subjected to or otherwise control further processing, for example to allow the rapid calculation of complex lighting effects.

Texture mapping can be implemented entirely in software, but in the context of the present invention we are concerned with hardware implementations in the field of real-time image synthesis. To avoid aliasing effects, it is necessary to filter the texel values during mapping. To simplify the computation of filtered values it has become conventional to store textures in so-called pyramidal arrays, comprising a succession of 2-D arrays, each pre-filtered to a different level of resolution. The generation and storage of pyramidal texture arrays are described by Lance Williams in a paper entitled "Pyramidal Parametrics" in Computer Graphics, Volume 17, No. 3 (Proc. SIGGRAPH 1983) at pages 1 to 11.

The interpolation means are provided to avoid sharp transitions in the displayed resolution as an object moves further away or closer to the viewpoint. However, because texel values must be read from two level arrays, this improvement doubles the number of memory read cycles that are required, unless parallel readout from both arrays is possible. Williams suggests a scheme whereby all levels of a given texture pyramid can be accessed simultaneously, but the number of connections that this entails makes the system very expensive and therefore unsuitable for mass market applications. Furthermore, the known schemes depend on a fixed "MIP map" arrangement of the arrays within a 2-D memory. The provision of a fully dual-ported texture memory would be another solution, but is also impractical because of the small capacity of dual-ported memory devices.

It is an object of the invention to allow parallel readout from two levels of a pyramidal array by a simpler and/or more economical technique than those previously considered.

The invention provides a display apparatus as set forth in the opening paragraph, characterised in that the texture memory comprises separate first and second parallel-addressable texture memory banks and in that the apparatus further comprises means for storing 2-D arrays of texel values forming successive levels of the stored pyramidal array alternately in only one of the first and second texture memory banks. This enables simultaneous read-out of texel values from two different levels of the pyramidal array for supply to the interpolation means, without excessive parallelism or duplication of storage space.

The invention is based on a recognition that the degree of parallelism required is limited, and can be provided by just two separate single-ported banks of texture memory, irrespective of the number of levels in the pyramidal array. All that is necessary is to ensure that adjacent levels of any given pyramid are always stored in different banks.

Each texture memory bank may be divided into at least three parallel-addressed memories, the texel value storage means being arranged to distribute the texel values being stored in an interleaved manner so that values for a 2-D patch of texels may be read in parallel from each of the first and second texture memory banks, the interpolation means being arranged to combine the inter-level interpolation with 2-D interpolation within each patch to generate a single 3-D interpolated value from a set of eight or more stored texel values. Thus full 3-D interpolation can be implemented with only one memory read cycle time per pixel.

The interpolation means may comprise first and second 2-D interpolators, for performing 2-D interpolation between the values within the patches stored in the first and second texture memory banks respectively, and a 1-D interpolator for combining interpolated values generated by the first and second 2-D interpolators so as to perform interpolation between the two adjacent levels of the array. The 2-D interpolators may be simple

bilinear interpolators and the 1-D interpolator may be a simple linear interpolator, making the interpolation means as a whole a trilinear interpolator.

In such an apparatus, the display processor may further comprise feedback means whereby texel values read from a 2-D array stored in one texture memory bank and interpolated by the first 2-D interpolator may be written into a smaller, corresponding array in an other texture memory bank, so that the smaller array represents the same pattern of modulation as the first-mentioned array but at a lower level of resolution. Such a feedback means enables the display processor to generate and store all or part of a pyramidal array of texel values from a single externally generated 2-D array. The fast 2-D interpolators can thus be put to a second use, in accordance with another invention disclosed herein eliminating the need for separate hardware or software filters to generate each pyramid and opening up the possibility of real-time generation of texture pyramids.

The display apparatus may comprise physical address generating means for receiving pyramidal texture coordinates in the form of a 2-D coordinate pair and an associated level coordinate, and for generating therefrom a physical address for application to the texture memory, the said means further comprising means for generating a further physical address corresponding to the same 2-D texture coordinate pair but with a level coordinate offset from the received level coordinate, and selection means whereby the physical address and the further physical address are applied to different ones of the first and second texture memory banks so as to cause the said simultaneous read-out of texel values from two adjacent levels of the pyramidal array.

Each physical address may comprise two physical address coordinates, each generated from a respective one of the received 2-D texture coordinate pair independently of the other address coordinate. It is conventional in texture mapping hardware to use a 2-D addressable memory. In particular this enables the use of the "multum in parvo" or "MIP map" method of storage as described by Williams which allows a compact storage and very efficient retrieval of pyramidal texture maps.

Alternatively, in accordance with another invention described herein, each physical address may comprise a linear address generated by combining both coordinates of the received 2-D texture coordinate pair. This other invention provides freedom to store each 2-D array in any suitable space in the memory, allowing more efficient memory utilisation than is possible in a 2-D memory.

Embodiments of the invention will now be described, by way of example, with reference to the accompanying drawings in which :

Figure 1 is a block diagram of a display apparatus including a known type of texture mapping hardware ;

Figure 2 shows a 2-D array of texel values stored in a texture memory ;

Figure 3 is a block schematic diagram of novel texture mapping hardware embodying the present invention;

Figure 4 illustrates the storage of three pyramidal texture maps in two linear texture memories in the hardware of Figure 3 ;

Figure 5 illustrates the contents of a page table memory in the hardware of Figure 3 ;

Figure 6 is a flowchart illustrating the generation of filtered 2-D arrays using feedback paths in the hardware of Figures 3 to 5 ; and

Figure 7 illustrates the storage of two pyramidal texture maps in two two-dimensional texture memories similar to that of the known hardware in accordance with the present invention.

Figure 1 is a block diagram of a display apparatus including a known type of texture mapping hardware. A keyboard 10 and a tracker ball-type input device 12 provide input from a user to a central processing unit (CPU) 14. The tracker ball may be used for designing 3-D objects to be manipulated by the system, in a known manner. Other input devices may also be used, of course, such as a joystick, digitising-tablet, or a "mouse". Such devices may also be used to manipulate images created by rotating, zooming etc. In general, such devices can be used more intuitively and efficiently than a conventional keyboard alone. Objects and also photographic images to be applied to object surfaces by texture mapping can also be input from a video source such as a camera 16.

The CPU 14 is connected via a bus 18 (for example a VME bus) to a disc store 20, a ROM 22 and a main memory (MRAM) 24. The disc store, which may include magnetic floppy discs, hard discs, and/or optical memory discs, is used for storing data (for example images or 3-D model data) which can then be recalled and manipulated to generate new images as desired. Such data may include the user's work from previous input sessions, and/or commercially generated data, for example for use in interactive computer-aided design or computer simulations for education or entertainment. To allow modelling of 3-D objects, such data will generally be stored as polygonal model data rather than in the form of two-dimensional images. In that case, the data corresponds to a 3-D model containing objects which are typically broken down into groups of polygonal surfaces (primitives) in a 3-D "object" space (triangular or quadrilateral surfaces for example). The data for each object in the model comprises a list giving the position and nature of every polygon that goes to make up the object, including the relative positions of its vertices and the colour or transparency of the polygon surface. In other systems, primitives may comprise curved surface patches, as is known in the art. It is known that a "text-

ure" can be specified for mapping onto the surface, so that detail can be represented without increasing the number of primitives that make up the scene. A texture map is a stored 2-D array of texture element ("texel") values defining a 2-D pattern of modulation that may for example define the colour of pixels in a manner to described below. The texture may alternatively modulate other quantities such as reflectance or surface normal direction, as is known in the art. These texture maps may also be stored in the disc store 20 and recalled as required.

The CPU 14 and the other components of the system then translate the 3-D model "world" in object space into a two-dimensional view for the user (in "viewer" space), from whatever viewpoint the user chooses, by means of geometric transformations effecting translations, rotations and perspective projections, generally by means matrix multiplication of vertex coordinates. The CPU 14 may also perform clipping and lighting calculations on a per-primitive or per-vertex basis.

The ROM 22 and MRAM 24 provide program memory and workspace for the CPU 14, which may comprise a microprocessor, such as a Motorola MC68020. Special processing hardware 26 may be provided to assist the CPU 14 to perform the large number of arithmetic operations required to convert all but the simplest models into a two-dimensional scene. The hardware 26 may comprise standard arithmetic circuits or it may include more powerful custom-built or programmable digital signal processing (DSP) integrated circuits, and may be connected to the CPU 14 for example via a VME bus connection. The nature of the hardware 26 will depend on the requirements of the system, for example with respect to speed, resolution, number of primitives per scene, etc.

A display processing unit (DPU) 28 is connected between outputs of the CPU 14 (the bus 18) and inputs of a display memory (VRAM) 30. The display memory 30 stores pixel data COL in raster-scan format. The pixel data COL might typically include for each pixel three 8-bit values (total 24 bits) corresponding to red (R) green (G) and blue (B) components of the desired image. Those skilled in the art will appreciate that in other embodiments fewer or more bits may be provided for, or the bits might define the colour in terms of different components.

In the DPU 28 the primitives are "scan converted" so that they may be drawn into the display memory 30. Scan conversion is a process whereby the pixels covered by each primitive are written row by row and pixel by pixel, in the same way that the complete image will be scanned for output to the display.

A timing unit (video controller) 32 generates read-address signals XD and YD to address the pixel data within the VRAM 30 synchronously with the raster-scanning of a display screen 34. In response to these address signals, the locations in the VRAM 30 are scanned row by row and column by column to read colour values COLD which are fed to a digital to analogue converter (DAC) 36. If a non-RGB colour code is used, a matrix circuit or colour look-up table may be provided to translate the pixel data COLD into the equivalent RGB signal for supply to the display screen 34, which may for example be a cathode-ray tube (CRT) display screen. The display 34, directly or indirectly, also receives timing signals (SYNC) from the timing unit 32.

To draw or "render" a primitive, the CPU 14 (or the special hardware 26) causes registers within the DPU 28 to be loaded, via the bus 18, with values defining a single primitive (for example in terms of vertex coordinates, edge slope and so on) and its various attributes - colour, reflectance and so forth. The DPU 28 then generates pixel coordinates (X and Y) so as to scan systematically the entire area covered by the primitive. The pixel coordinates X and Y are applied as write addresses to the VRAM 30, so that a pixel value COL can be written into the VRAM 30 for every pixel.

The pixel values COL can be generated so that a basic surface colour of the primitive is modulated to account realistically for attributes of an object's surface (for example colour, transparency, diffuse reflectance, specular reflectance) and of the 3-D environment (for example locations, colours and shapes of light sources, distance haze). This modulation can be generated arithmetically from parameters loaded with the primitive data, for example to produce smoothly varying shading to simulate a curved surface. However, to provide more detailed modulation, it is known to use mapping hardware such as that referenced 40 to supply modulation values MOD according to a predetermined pattern stored in advance in a texture memory 41.

To this end, the DPU 28 generates a pair of texture coordinates U and V simultaneously with each pair of pixel (display) coordinates X and Y so that the modulation pattern is mapped onto the primitive surface, implementing geometric transformations (i) from texture space into object space and (ii) from object space into viewer (display) space. Figure 2 provides an illustration of the relationship between texture space, defined by the horizontal and vertical axes labelled U and V, and screen space defined by the oblique axes X and Y. The actual stored texel values correspond to integer values of U and V and are represented by a square array of solid circular dots. The locations of the pixels in screen space are marked by diagonal crosses ('x') and lie along scanlines referenced S1, S2 and S3 etc. parallel to the X-axis.

To define the coordinates U and V required to address texel values corresponding to the series of pixel values on the scanlines S1, S2, S3 etc., the CPU 14 (or drawing hardware 26) may for example provide the

DPU 28 in advance with the coordinate pair (U_0, V_0) corresponding to the first pixel on scanline S1, and also partial derivatives $\partial U/\partial X$ and $\partial V/\partial X$ defining the slope of the screen space scanlines S1 etc. in texture space and partial derivatives $\partial U/\partial Y$ and $\partial V/\partial Y$ defining the slope of the pixel columns in texture space. In the example illustrated, the transformation from texture space to screen space is linear. In a more general case, the scanlines S1 etc. and the pixel columns might diverge or converge, or even curve, in which case the partial derivatives vary from point to point across the primitive.

The texture coordinates U and V are processed within the mapping hardware 40 in a manner to be described below and applied to the texture memory 41 so that a modulation value MOD is available for each pixel location X, Y being addressed. The value MOD commonly comprises a colour value, and in principle it could directly form the pixel value COL and be fed directly into the display memory (VRAM) 30, as shown by the dotted data path 42. More commonly, however, even if the values MOD are colour values, they will require to be modified within the DPU 28 to allow for realistic lighting effects. In a more general case, the modulation values MOD are used within the DPU 28 together with other parameters to modify the pixel values COL less directly. For example, in so-called "bump mapping", the values MOD modulate the surface normal direction of the primitive, so as to affect subsequent lighting calculations and so, indirectly, the pixel values COL. Another technique, known as "environment mapping" uses the TRAM to store an image of the environment, for example using U and V as spherical coordinates, so that specular reflections of a complex environment (including light sources, windows, other objects and so on) can be simulated. These and various other applications of mapping hardware are summarised in an article "Survey of Texture Mapping" by Paul S. Heckbert in IEEE Computer Graphics and Applications, November 1986 at pages 56 to 67. Those skilled in the art will recognise that the invention may be applied in all such applications of mapping hardware.

It is known that the texels represented in the texture memory 41 will not in general correspond on a one-to-one basis with the pixels of the display and, in particular when the primitive is shown in the distance and the texture is consequently mapped onto a very small number of pixels, two-dimensional spatial filtering is required to avoid the aliasing effects that would be disturbing to the viewer if simple sub-sampling were used.

It is further known that a generalised filter cannot be applied economically in an apparatus where real-time moving images are to be synthesised, and the reference Williams describes the conventional solution to this which is to store several 2-D arrays (hereinafter referred to as "maps") for a given pattern, each being successively smaller and pre-filtered to a successively lower resolution. The DPU 28 then need only produce a level coordinate L to determine the appropriate map to use. For compact storage and for high speed access to the texel values, the maps may be chosen to be square, having power-of-two dimensions, and be stored in a square texture memory according to the "multum in parvo" ("MIP map") technique described by Williams.

Figure 1 shows within the texture memory 41 the colour components R, G and B of a texture pyramid stored as a MIP map. The largest (highest resolution) map ($L=0$) may for example comprise 512×512 texels, the $L=1$ maps comprise 256×256 texels and so on down to $L=9$ where each map becomes a single texel. Assuming, for the sake of example, that each texel value comprises an 8-bit value for each of the R, G and B colour components, the entire texture memory 41 is thus 1 Mbyte in size.

The texel values are stored in the memory 41 in advance of rendering by the CPU 14 via the bus 18 and a writing port 43 of the memory 41. For each texel value to be read, the DPU 28 generates a 2-D coordinate pair, each coordinate (U and V) of which includes at least an integer part 9 bits in length. At the same time, the level coordinate L is generated by the DPU 28 and used to generate physical coordinates U' and V' from the "virtual" coordinates U and V for application to read address ports 44 and 45 respectively of the texture memory 41. In response to each physical coordinate pair U' , V' , the memory 41 releases the R, G and B components of an addressed texel via a (24-bit) read port 46.

Because of the two-dimensional binary tree arrangement of the MIP maps in the memory 41, the required physical coordinates U' and V' can be generated simply by a pair of binary shifting circuits 47 and 48 respectively, each right-shifting the respective coordinate a number of places defined by the level coordinate L. In particular, if $L=0$ represents the highest level, then the address corresponding to a given texel in the level 0 map can be converted to the physical address of the corresponding texel in the level L map can be found by right-shifting the U and V coordinates L places, effectively scaling-down each coordinate by 2^L . The level coordinate L can be supplied to the DPU 28 as part of the primitive data, but if perspective is to be accounted for in the mapping, then the level coordinate L will more probably be generated within the DPU on a per-pixel basis, dependent on the partial derivatives of U, V with respect to X, Y.

In order to allow full antialiasing, it is known to apply 3-D (for example, trilinear) interpolation between texel values, in which case the coordinates L, U' and V' can have fractional parts (L_f , U'_f and V'_f) as well as integer parts (L_i , U'_i , V'_i). The fractional parts of the U' and V' coordinates can be used to perform 2-D (for example, bilinear) interpolation between a square patch of four adjacent texels within one level, and the fractional part L_f of the level coordinate can be used to interpolate between (2-D interpolated) texel values from two adjacent

levels of the pyramidal array. To this end, it is necessary to read four texel values (U_i, V_i) , (U_i+1, V_i) , (U_i, V_i+1) and (U_i+1, V_i+1) from the level L_i map and four from the level L_i+1 map. Clearly a speed penalty is involved if these eight texel values are read serially. Fortunately, the four texel values for each level can be read in parallel via the read port 46 if the texture memory is constructed as four parallel memories interleaved to allow 2x2 patch addressing, as described hereinafter, enabling the eight values to be read in only two memory read cycles. It would be desirable, however, to enable both sets of four (L_i and L_i+1) texel values to be read in parallel and the Williams reference suggests that a hardwired addressing scheme could enable parallel access to all levels of a given MIP map. While this is possible in theory, the number of connections involved in the scheme proposed by Williams is too great to make it an economic solution for mass market applications. For example, with ten levels, 2x2 patch addressing (except at the lowest level) and 8 bits each for R, G and B per texel, 888 bits of data would need to emerge from the read port 46 of the texture memory 41 for every coordinate pair U,V applied.

In general, it will be desirable to store different texture pyramids in the texture memory 41. For example three texture pyramids might define the shapes 'o', '+' and 'x' mapped onto the faces of the cube shown on the screen of the display 34 in Figure 1. For this purpose, it is known to divide the square array at each level of the MIP map and store a mosaic of the corresponding 2-D arrays defining each 2-D pattern. The coordinate pairs U,V generated by the DPU 28 would then incorporate 2-D offsets to ensure that the correct part of the 2-D array is addressed. In this known technique, however, some of the space in the texture inevitably remains unused, effectively wasted. It is not possible in a general case to eliminate unused space from a mosaic of 2-D shapes. For example arranging the three square arrays representing the texture problems 'o', '+' and 'x' into the square array at each level in the texture memory 41 of the known hardware would result in at least one quarter of the available memory being wasted. Finding even an optimum solution to a general 2-D "jigsaw puzzle" is difficult, and would be quite impractical in real-time if arrays were allowed to take different shapes, such as squares, rectangles, triangles.

It is a further disadvantage of the MIP map approach that each texture pyramid occupies space at all levels (1 Mbyte in the example given above), even though, in a scene where the primitive is seen only in the distance, only one or two of the smaller maps ($L=5$, $L=6$ etc) may actually be being used at the time. It would be very advantageous if only the levels likely to be needed were stored in the texture memory at any one time and the freed space could be used for other texture maps. For example, even though the largest map ($L=0$) of the pyramid may never actually be read in the course of rendering an image, it still occupies three-quarters of the total texture memory storage. Unfortunately, it would be very difficult to provide the known hardware with the flexibility to overcome either of the above disadvantages.

Figure 3 shows novel mapping hardware that can be substituted for that shown at 40 in Figure 1. A linearly-addressed texture memory 41' is provided, so that the problem of eliminating wasted space is readily solvable. A texture management circuit 49 keeps track of the various arrays within the linear memory 41' and serves to convert the pyramidal coordinates L, U and V into linear physical texel addresses. Instead of using a 2-D offset to identify different textures all stored as a mosaic within a large map, the CPU 14 in the novel system supplies with the primitive data a texture identifying value T separate from the coordinates U and V. Any 2-D map forming part of a texture pyramid can thus be identified as map T.Li, where Li is the integer part of the level coordinate L. The circuit 49 is more complex than the simple 2-D MIP addressing hardware, but the improvement in memory utilisation and flexibility may be very great.

The DPU 28' in Figure 3 is a slightly modified version of the conventional DPU 28 (Figure 1), having an output T for passing the received identifying value to the texture management circuit 49. The modified DPU 28' also has an output carrying a logic signal FB for activating feedback paths as described in a later part of this description.

The novel hardware shown in Figure 3 also incorporates not only 2x2 patch addressing (to allow high-speed bilinear interpolation) but also a novel parallel structure so that eight texel values are available for trilinear interpolation simultaneously, yet without the excessive parallelism of the solution proposed by Williams. For this purpose, the texture memory 41' is divided into two banks of memory, TRAM1 and TRAM2, and the system ensures that arrays T.Li and T.Li+1 for two adjacent levels of a given texture pyramid will always be stored in different banks TRAM1 and TRAM2.

The texture management circuit 49 has inputs to receive the signals T, L, U and V from the DPU 28'. The texture management circuit includes a page location and logic circuit (PLLC) 50 which receives the texture identification T and at least the integer part Li of the level coordinate L supplied by the DPU.

The PLLC 50 stores information defining for each map T.Li (i) in which bank TRAM1 or TRAM2 the map T.Li is stored, (ii) the width w(T.Li) of the map in the U-direction and (iii) a base address B(T.Li) locating the start of the linearly-stored array in the appropriate bank TRAM1 or TRAM2.

In general terms, the PLLC 50 supplies the stored data for the maps T.Li and T.Li+1 to the remainder of

the circuit 49 which is thereby enabled to generate linear addresses A1 and A2 for application to the memory banks TRAM1 and TRAM2 respectively, so as to address the texel data for coordinate pair (U,V) in the levels Li and Li+1 in the texture pyramid T. The general formulae for these linear addresses are:

$$A(T.Li) = \frac{U}{su(T.Li)} + w(T.Li) \frac{V}{sv(T.Li)} + B(T.Li)$$

and

$$A(T.Li+1) = \frac{U}{su(T.Li+1)} + w(T.Li+1) \frac{V}{sv(T.Li+1)} + B(T.Li+1)$$

with address A1 = A(T.Li) or A(T.Li+1), depending on which of the two maps T.Li and T.Li+1 respectively is stored in memory bank TRAM1, and address A2 = A(T.Li+1) or A(T.Li) correspondingly. In the above expressions, su(T.Li) and sv(T.Li) represent generalised scale factors relating the dimensions of the map T.Li to those of the largest map (T.0) in the pyramid T, in the U- and V-directions respectively.

As in the known system, the addressing hardware in accordance with the invention can be greatly simplified if the scaling factors su(T.Li) and sv(T.Li) are limited to powers of two, defined for example by the expression $su = sv = 2^L$ for all values of L and T. Further simplification of the hardware can be obtained as in the MIP map system by limiting the width values w of maps to be powers of two texels, defined for example by a width index W(T.Li) and the expression $w(T.Li) = 2^{W(T.Li)}$. Both of these simplifying features are incorporated in the texture management circuit 49 shown in Figure 3. These restrictions allow not only square but also rectangular maps, but a hardware implementation of the more general formulae given above could be constructed if desired. The general formulae could also be further generalised to allow efficient storage of other shapes, for example triangular or trapezoidal textures, by causing the width w in the U-direction to vary across the map in the V-direction in a linear or even non-linear manner.

Taking into account these limitations introduced to simplify the hardware, new formulae can be derived for translation into simplified hardware. These formulae are given below and use the symbols " \rightarrow " and " \leftarrow " to indicate a binary right shift (divide) and left shift (multiply) respectively. Thus for example the expression " $(U \rightarrow L1)^i$ " indicates the integer part of a value U after right-shifting by L1 bit positions, in other words the quotient of U and 2^{L1} .

$$A1 = (U \rightarrow L1)^i + ((V \rightarrow L1)^i \leftarrow W1) + B1$$

where $L1 = Li$ or $Li+1$ and $w(T.L1) = 2^{W1}$, and

$$A2 = (U \rightarrow L2)^i + ((V \rightarrow L2)^i \leftarrow W2) + B2$$

where $L2 = Li+1$ or Li correspondingly, and $w(T.L2) = 2^{W2}$.

Returning to Figure 3 which shows a hardware implementation of these formulae for A1 and A2, the PLLC 50 in Figure 3 has outputs supplying values W1, B1, W2 and B2 defining the widths and base locations for the maps to be addressed in TRAM1 and TRAM2 respectively. The PLLC 50 also generates a binary signal SWI1 which takes the value 1 or 0 depending on whether texture memory bank TRAM1 contains the map T.Li or T.Li+1 respectively and a complementary signal SWI2 = $\overline{SWI1}$ which gives the corresponding indication in relation to the other bank TRAM2 of the texture memory 41'.

In the remainder of the texture management circuit 49, an adder 51 is provided to generate the value Li+1 from the value Li generated by the DPU 28'. A multiplexer 52 responsive to the logic signal SWI1 selects either Li or Li+1 to generate the level coordinate L1 for the map stored in bank TRAM1. Another multiplexer 53 is responsive to the complementary logic signal SWI2 selects the other of Li and Li+1 to generate the level value L2 for the map stored in bank TRAM2. A first right-shifter 54 responsive to the level coordinate L1 receives the U coordinate of the pair U,V generated by the DPU 28' and generates a first scaled U coordinate $U1 = U \rightarrow L1$. A second right-shifter 55 responsive to the same level coordinate L1 receives the V coordinate and generates a first scaled V coordinate $V1 = V \rightarrow L1$. Third and fourth right-shifters 56 and 57 also receive the U and V coordinates respectively and are responsive to the level coordinate L2 to generate second scaled U and V coordinates $U2 = U \rightarrow L2$ and $V2 = V \rightarrow L2$ respectively for the map stored in the second bank TRAM2 of the texture memory 41'.

The scaled coordinates U1, V1, U2 and V2 are all separated into their integer parts U1i etc. and their fractional parts U1f etc. First and second left-shifters 58 and 59 receive the integer parts V1i and V2i of the scaled V coordinates V1 and V2 respectively and are responsive to the width indices W1 and W2 respectively so as to generate values $2^{W1}.V1i$ and $2^{W2}.V2i$ respectively, where 2^{W1} and 2^{W2} are the widths of the maps stored in the banks TRAM1 and TRAM2 respectively.

An adder 60 adds the integer part of $U1i$ of the first scaled U coordinate $U1$ to the value $2^{W1}.V1i$ generated by the first left-shifter 58 to generate a first linear offset address $I1$. A further adder 61 adds the first offset address $I1$ to the first map base address $B1$ generated by the PLLC 50 to generate the first linear texel address $A1$ for application to the first bank $TRAM1$ of the texture memory 41'. Similarly, a further adder 62 generates a second linear offset address $I2$ by adding the values $U2i$ and $2^{W2}.V2i$, while a still further adder 63 adds the second linear offset address $I2$ to the second map base address $B2$ generated by the PLLC 50 to generate the second linear texel address $A2$ for application to the second bank $TRAM2$ of the texture memory 41'. Since the bits of value $V1i$ (or $V2i$), once shifted, do not overlap with those of the value $U1i$ ($U2i$), the adders 60 and 62 can in fact be implemented by simpler OR-gates.

Each texture memory bank $TRAM1$ and $TRAM2$ is further divided as shown into four parts A, B, C, and D which can be addressed in parallel. The texel values defining a given map are distributed between the four parts A-D of the appropriate memory bank ($TRAM1$ or $TRAM2$) according to a predetermined pattern such as that illustrated by the letter A, B, C or D next to each texel value (solid circle) in Figure 2, so as to allow parallel addressing of a 2×2 patch of texels. In the example pattern shown, in an even numbered line of texels (V even), texel values are stored alternately in parts A and B. In odd-numbered lines (V odd) the values are stored alternately in parts C and D.

To enable this patch addressing a special address port 64 receives the linear texel address $A1$ from the output of adder 61 and generates therefrom four addresses $A1A-A1D$ for application to the four memories $TRAM1A$ to $TRAM1D$ respectively, in response to which the texel values for the patch (U,V) , $(U+1,V)$, $(U,V+1)$ and $(U+1,V+1)$ become available via four corresponding read ports 65A-65D.

To enable generation of the correct four addresses $A1A-A1D$, the address port 64 receives the least significant bits $U1lsb$ and $V1lsb$ of the integer parts of the first scaled coordinate pair $U1,V1$, which define whether $U1$ and $V1$ respectively are odd or even. With regard to the detailed design of the patch addressing hardware 64, 65A-65D, this can be similar to that used for transforming digitised video images, an example of which is shown in Figure 2 of an article "Transforming Digital Images in Real Time" by Joel H. Dedrick in ESD : The Electronic System Design Magazine, August 1987 at pages 81 to 85. One difference from the known hardware is necessitated by the linear storage of the texel arrays in the system shown in Figure 3. In the Dedrick circuit, which uses a 2-D addressable framestore memory, a unit value 0001 is added to the vertical coordinate (Y') to address the texel values in the next row ($Y'+1$) of the image. In the circuit of Figure 3, however, the width $w(T.L) = 2^{W1}$ of the array must be added to the linear address $A1$ to address correctly the texel values for the next row $V1+1$ of the linearly stored texture map. To this end, the address port 64 also receives the width index $W1$ generated by the PLLC 50 for the map stored in bank $TRAM1$.

A similar patch address port 66 is provided for the second bank $TRAM2$ of the texture memory 41' and receives the second linear address $A2$, odd/even indicators $U2lsb$ and $V2lsb$ and the second width index $W2$. The port 66 generates patch addresses $A2A-A2D$ which are applied to respective parts A-D of the second texture memory bank $TRAM2$ which has four corresponding read ports 67A-67D.

It may be noted that many alternative arrangements may be suitable for generating the patch addresses $A1A-A1D$ and $A2A-A2D$. For example, instead of generating the single linear address $A1$ and then expanding it to form addresses $A1A-A1D$, it may be advantageous to integrate the patch addressing function with the linear address generating function, to generate each of the addresses $A1A-A1D$ directly from the coordinates $L1$, $U1$ and $V1$. While some components may need to be quadruplicated in such an embodiment, other components can contribute to the generation of at least two of the addresses $A1A-A1D$. It will also be appreciated that larger patches could be addressed with more parallel memories and suitable ports.

The four texel values from the read ports 65A-65D of bank $TRAM1$ are supplied to inputs of a first bilinear (2-D) interpolator $BIL1$ which also receives the fractional parts $U1f$ and $V1f$ of the first scaled coordinate pair $U1,V1$. The bilinear interpolator $BIL1$ combines the four texel values in the patch addressed by addresses $A1A-A1D$ (derived from the integer parts $U1i$ and $V1i$ of the pair) so as to generate a first bilinearly interpolated texel value $MOD1$. The texel values from the read ports 67A-67D of bank $TRAM2$ are similarly applied to a second bilinear interpolator $BIL2$ which also receives the fractional parts $U2f$ and $V2f$ of the second scaled coordinate pair $U2,V2$ and generates a second bilinearly interpolated texel value $MOD2$.

The two bilinearly interpolated values $MOD1$ and $MOD2$, one derived from the map $T.L$ and the other derived from the map $T.L+1$ are then fed to a linear interpolator $LINT$. The interpolator $LINT$ combines the values $MOD1$ and $MOD2$ in proportions determined by the fractional part Lf of the level coordinate L received from the DPU 28 to generate a trilinearly interpolated modulation value MOD for the pyramidal coordinates L , U and V . As in the known apparatus (Figure 1), the value MOD may be used to effect a modulation of pixel colour values COL either directly (dotted path 42) or indirectly via further processing in the DPU 28.

Where the texel values MOD define colours, it will be appreciated that each texel value will comprise three colour component values such as R, G and B, and the interpolators $BIL1$, $BIL2$ and $LINT$ may in fact comprise

three interpolators each, or may be adapted in some other way to perform the three-component interpolation.

The hardware shown in Figure 3 also incorporates first and second feedback paths 70 and 72 so that the bilinearly interpolated values MOD1 from the first memory bank TRAM1 can be fed into a write port 71 of the second bank TRAM2 of the texture memory 41' and the values MOD2 from the second bank TRAM2 can be fed into a write port 73 of the first bank TRAM1. The logic signal FB supplied by the DPU 28' indicates in the '1' state that a feedback path is to be activated. AND gates 74 and 75 combine the signal FB with the logic signals SWI1 and SWI2 respectively generated by the PLLC 50 to generate a pair of logic signals FB1 and FB2, respectively. Two multiplexers 76 and 78 are responsive to the signals FB1=1 and FB2=1 respectively to complete either the first or second feedback path 70 or 72 respectively when FB AND SWI1 = 1 or FB AND SWI2 = 1 respectively. At all other times, each multiplexer 76 or 78 serves to connect the corresponding write ports 71 or 73 with the CPU 14 via the bus 18. The purpose of the feedback paths 70 and 72 will be described in due course but for the moment it should be assumed that they are not active (FB=0).

Figure 4 shows, by way of example, three part-pyramidal texture maps ($T = 1, 2, 3$) stored in the linear memory banks TRAM1 and TRAM2 so as to allow parallel access to all the texel values required for trilinear interpolation. Each memory bank constitutes a linear physical address space A1 or A2 respectively. The 2-D array (map) of texel values for level Li of map T is stored in linear form as a page of data referenced PT.Li. Texture 1 can be seen to have four pages P1.0, P1.1, P1.2 and P1.3, each one quarter of the linear size of the previous page. While P1.0 is stored in bank TRAM1, page P1.1 is stored in bank TRAM2, P1.2 in bank TRAM1 and so on, alternating for as many levels as are stored.

The second texture ($T=2$) is stored as three pages P2.1 to P2.3, alternating between banks TRAM1 and TRAM2, each with a corresponding entry in the page table memory PTAB, described below with reference to Figure 5. A highest level page P2.0 exists, but, because it is not required at the time illustrated, it is not stored in the texture memory, so that more free space (shaded) is available to receive other maps as required. In fact, the memory map shown in Figure 4 may represent a time when many different maps, forming various levels of various texture pyramids, have been loaded, used and then deleted as the corresponding texture has become not required, or not required at such a high (or low) resolution. It should be noted that even if more texel data is required than there is room for in the banks TRAM1 and TRAM2 of the texture memory 41', at least a low resolution map can be loaded and interpolated to the size required until the space does become available. This "graceful degradation" characteristic of the linear texture memory contrasts favourably with the known 2-D MIP map memory in which a texture pyramid generally occupies space at all levels or not at all.

The third texture ($T=3$) is also only loaded in three pages, again alternately between TRAM1 and TRAM2, but page P3.0 is smaller than the Li=0 pages of the first two textures, simply because there is less detail required to define the third texture. This again allows more efficient usage of texture memory capacity compared with the conventional MIP map.

Figure 5 shows the contents of the page table memory PTAB, shown dotted within the page location and logic circuit PLLC 50 of Figure 3, which enables a given array T.Li to be found in the memory. There is an entry in the page table for each page PT.Li of each texture. A first item in each entry is a single bit value SWI1/SWI2 which takes the value 1 or 0 depending on whether the page is stored in the bank TRAM1 or the bank TRAM2 respectively. This bit value therefore enables the logic circuit PLLC to generate the signals SWI1 and SWI2 to control the generation of the addresses A1 and A2 as described with reference to Figure 3.

A second item in each entry in the table is the width index W of the map T.Li. The page P1.0 contains a map $2^W = 2^9 = 512$ texels wide, P1.1 a map $2^8 = 256$ texels wide, P3.1 a map $2^7 = 128$ texels wide and so on. A third item in each entry contains the base location B in the memory (TRAM1 or TRAM2) starting at which the texel values comprising that page are stored.

Returning to Figure 3, the arrays of texel values forming the pages P1.0 etc. can be stored in the memory banks TRAM1 and TRAM2 by the CPU 14 via the bus 18 and write ports 71 and 73 respectively as shown. It is a straightforward matter of "housekeeping" for the CPU 14 to ensure that (i) any new page of texel data is stored in an otherwise unused part of the texture memories, (ii) each alternate level of a given pyramid is stored in a different bank TRAM1 or TRAM2 and (iii) that the page table PTAB in the PLLC 50 is at the same time loaded with the appropriate values of SWI1/SWI2, W and B. If enough space is not available in a single block for a new texture (perhaps P2.0 is to be loaded), then a "garbage collection" operation can be carried out to gather the unused areas together into a large enough area.

Entire pyramidal texture arrays may be stored in a database in the disc store 20 or in the main memory 24, with the various 2-D level arrays (maps) being transferred to spare locations in the appropriate texture memory as required. However, the provision of the feedback paths 70 and 72 enables the bilinear interpolators BIL1 or BIL2 to be used as filters in a manner to be described below to generate maps for successive levels from a single, high resolution map loaded via the bus 18. This may be advantageous, since the hardware with patch-addressing and interpolation may provide a much faster filter than the conventional alternative which is a

software routine executed by the CPU 14.

Figure 6 shows a flowchart defining a sequence of operations performed to generate a filtered map T.Li+1 from a map T.Li previously stored in one bank of the texture memory 41', using the feedback of values generated by the existing bilinear interpolators. In a first step 80, the CPU 14 allocates texture memory space for the page PT.Li+1 by loading an entry comprising appropriate values SWI1/SWI2 (T.Li+1), W(T.Li+1) and B(T.Li+1) in the page table PTAB within the PLLC 50.

Next, in a step 82, the DPU 28' is caused to set the logic signal FB = 1 to enable activation of the appropriate feedback path 70 or 72. In step 84, the DPU 28' arranges that any values MOD generated by the linear interpolator LINT during the filtering process are ignored. In a step 86, the DPU 28 is set up as it would be for drawing a polygon, with values for U_0 , V_0 and the partial derivatives chosen so as to cause the generation of the desired filtered values. Referring to Figure 2, suitable texel positions for the filtered map T.Li+1 are marked by circled crosses ("x"). To cause the generation of values interpolated to these texel positions, the starting position would be $(U_0, V_0) = (\frac{1}{2}, \frac{1}{2})$, with partial derivatives as follows:

$$\frac{\partial U}{\partial X} = 2; \quad \frac{\partial V}{\partial X} = 0; \quad \frac{\partial U}{\partial Y} = 0; \quad \frac{\partial V}{\partial Y} = 2.$$

In step 88, the DPU 28' is caused to "draw" the imaginary polygon, leading to the automatic generation of the desired filtered values MOD1 or MOD2 at the output of the bilinear interpolator BIL1 or BIL2, depending on whether the source map T.Li is stored in bank TRAM1 or TRAM2. At the same time, because logic signal FB = 1 and either SWI1 = 1 or SWI2 = 1, either FB1 = 1 or FB2 = 1, and consequently one of the multiplexers 76 and 78 is activated to complete the appropriate feedback path 70 or 72. The interpolated values are then automatically written into the locations in the other memory bank (TRAM2 or TRAM1) that were allocated for the map T.Li+1 in step 80.

The process of Figure 6 can be repeated if required, incrementing the level coordinate L each time as shown dotted in step 90, to use the feedback paths 70 and 72 alternately, until an entire pyramidal array may have been generated within the display processor from a single externally generated high-resolution array. Those skilled in the art will appreciate that different scaling factors are possible, and asymmetrical filtering would be possible in embodiments where asymmetrical maps are allowed for, for example to generate a rectangular map from a square one. As a particular example, using

$$(U_0, V_0) = (0, 0) \text{ with } \frac{\partial U}{\partial X} = 1; \quad \frac{\partial V}{\partial X} = 0; \quad \frac{\partial U}{\partial Y} = 0 \text{ and } \frac{\partial V}{\partial Y} = 1$$

would provide a texel-for-texel transfer of a map from one bank of the texture memory to the other. Such a transfer may be useful in the "garbage collection" operations mentioned above.

The use of the texture mapping hardware itself for map filtering and/or transfer has advantages over the conventional method using software running on the host computer, because the display processor hardware is already specialised for rapid addressing of texture arrays, including 2 x 2 patch addressing hardware, interpolating hardware geared to receive R, G and B in parallel, and so forth. This advantage can be readily appreciated by considering that filtering or transferring a 2-D array in this apparatus takes about the same amount of time as it would take to render a single, similarly sized polygon. This is an overhead that can generally be absorbed in real-time without a significant loss in performance since a typical system may already be rendering hundreds or even thousands of polygons in each image frame. The use of feedback to implement a hardware texture filter is not dependent on the linear nature of the map storage, nor on the dual-bank memory construction. Even in a conventional 2-D MIP map arrangement such Figure 1, and even if parallel access to the two patches of texels were not provided, the use of the specialised addressing and interpolating hardware to generate filtered maps still offers a great speed advantage over the conventional software approach, and the overhead is still only equivalent to one polygon.

The limitations on map sizes and so forth imposed in the embodiment described simplify the construction, for example, of the PLLC 50 and allow the use of bit shifters 54 to 59 instead of complex multipliers. These limitations also simplify the allocation of space to new pages in the texture memory. Clearly, however, the design need not be limited to this case, and some variations will be enumerated below. For example, the width index W would not need to be stored separately if it were known that a map's width was always 2^{N+1} , where 2^N is the size of the largest permitted map. Thus, in the example of Figures 4 and 5 where the highest level of the texture T=3 is only 256 texels wide, the width index could be eliminated if the page P3.0 were renamed P3.1, and so

on, since it could then be known that all maps with $L_i=1$ have $W=8$ (width=256), whether or not there exists a larger map with $L_i=0$. To cater for asymmetrically filtered maps, two values L_u and L_v could be supplied, identifying a previously stored map of 2^{N-L_u} texels by 2^{N-L_v} rows. Other variations that may be allowed, will be readily apparent to those skilled in the art who will also readily appreciate the variations in the structure of the texture management circuit that would be required to implement such variations.

Those skilled in the art will further appreciate that the principle of providing a dual-bank texture memory to allow parallel access for inter-level interpolation is not only applicable to the linear texture memory arrangements described above. Figure 7 illustrates one possible way of storing two pyramidal texture arrays (R1/G1/B1 and R2/G2/B2) in two two-dimensional texture memory banks 170 and 172. The allocation is similar to the conventional MIP map arrangement of Figure 2 but allows parallel read-out from adjacent levels of any one texture. Those skilled in the art will readily appreciate how the mapping hardware 40 of Figure 1 can be adapted to incorporate and take advantage of a dual texture memory in accordance with this scheme, to achieve parallel addressing of the two level maps required for inter-level interpolation. The provision of feedback paths analogous to the paths 70 and 72 (Figure 3) would also allow the hardware to generate pyramids from single high-resolution maps.

From reading the present disclosure, yet other variations will be apparent to persons skilled in the art. Such variations may involve other features which are already known in the design, manufacture and use of electronic graphics systems, texture mapping hardware and component parts thereof and which may be used instead of or in addition to features already described herein. Although claims have been formulated in this application to particular combinations of features, it should be understood that the scope of the disclosure of the present application also includes any novel feature or any novel combination of features disclosed herein either explicitly or implicitly or any generalisation thereof, whether or not it relates to the same invention as presently claimed in any claim and whether or not it mitigates any or all of the same technical problems as does the present invention. The applicants hereby give notice that new claims may be formulated to such features and/or combinations of such features during the prosecution of the present application or of any further application derived therefrom.

Claims

1. A display apparatus comprising a display processor having associated display memory and texture memory for storing at least one pyramidal or part-pyramidal array of texture element ("texel") values representing a two-dimensional (2-D) modulation pattern at at least two distinct levels of resolution identifiable by respective values of a level coordinate, the apparatus further comprising inter-level interpolating means responsive to a fractional part of a received level coordinate for combining together corresponding texel values from two levels of a stored pyramidal array so as to generate an interpolated texel value, characterised in that the texture memory comprises separate first and second parallel-addressable texture memory banks and in that the apparatus further comprises means for storing 2-D arrays of texel values forming successive levels of the stored pyramidal array alternately in only one of the first and second texture memory banks.
2. A display apparatus as claimed in Claim 1 wherein each texture memory bank is divided into at least three parallel-addressed memories and wherein the texel value storage means are arranged to distribute the texel values being stored in an interleaved manner so that values for a 2-D patch of texels may be read in parallel from each of the first and second texture memory banks, the interpolation means being arranged to combine the inter-level interpolation with 2-D interpolation within each patch to generate a single 3-D interpolated value from a set of eight or more stored texel values.
3. A display apparatus as claimed in Claim 2 wherein the interpolation means comprise first and second 2-D interpolators, for performing 2-D interpolation between the values within the 2-D patches stored in the first and second texture memory banks respectively, and a 1-D interpolator for combining interpolated values generated by the first and second 2-D interpolators so as to perform interpolation between the two adjacent levels of the array.
4. A display apparatus as claimed in Claim 3 wherein the display processor further comprises feedback means whereby texel values read from a 2-D array stored one texture memory bank and interpolated by the first 2-D interpolator may be written into a smaller, corresponding array in another texture memory bank, so that the smaller array represents the same pattern of modulation as the first-mentioned array but at a lower level of resolution.

5. A display apparatus as claimed in any of Claims 1 to 4 comprising physical address generating means for receiving pyramidal texture coordinates in the form of a 2-D coordinate pair and an associated level coordinate, and for generating therefrom a physical address for application to the texture memory, the said means further comprising means for generating a further physical address corresponding to the same 2-D texture coordinate pair but with a level coordinate offset from the received level coordinate, and selection means whereby the physical address and the further physical address are applied to different ones of the first and second texture memory banks so as to cause the said simultaneous read-out of texel values from two adjacent levels of the pyramidal array.
6. A display apparatus as claimed in Claim 5 wherein each physical address comprises two physical address coordinates, each generated from a respective one of the received texture coordinate pair independently of the other address coordinate.
7. A display apparatus as claimed in Claim 5 wherein each physical address comprises a linear address generated by combining both coordinates of the received 2-D texture coordinate pair.

20

25

30

35

40

45

50

55

Fig. 1.

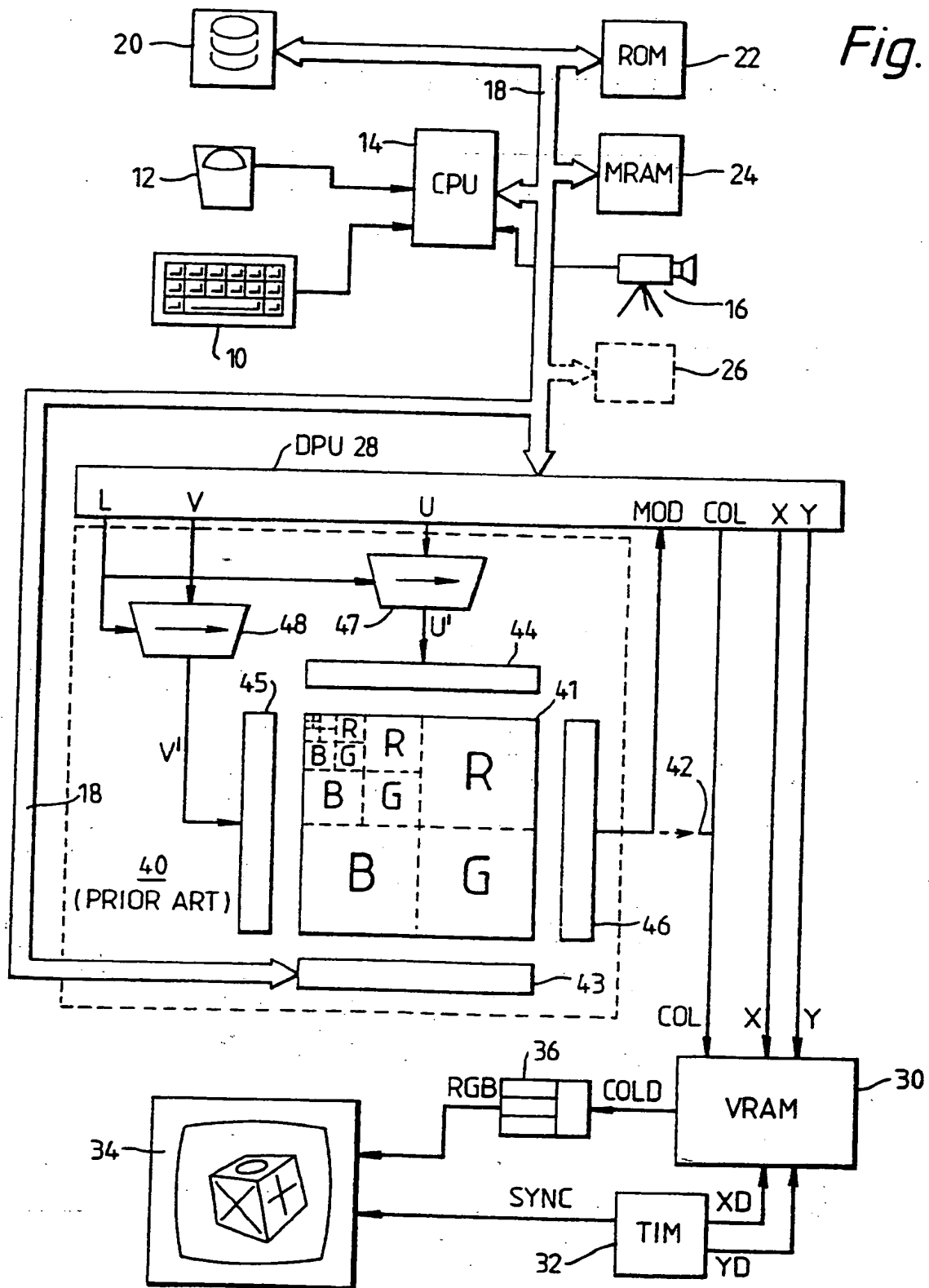


Fig. 2.

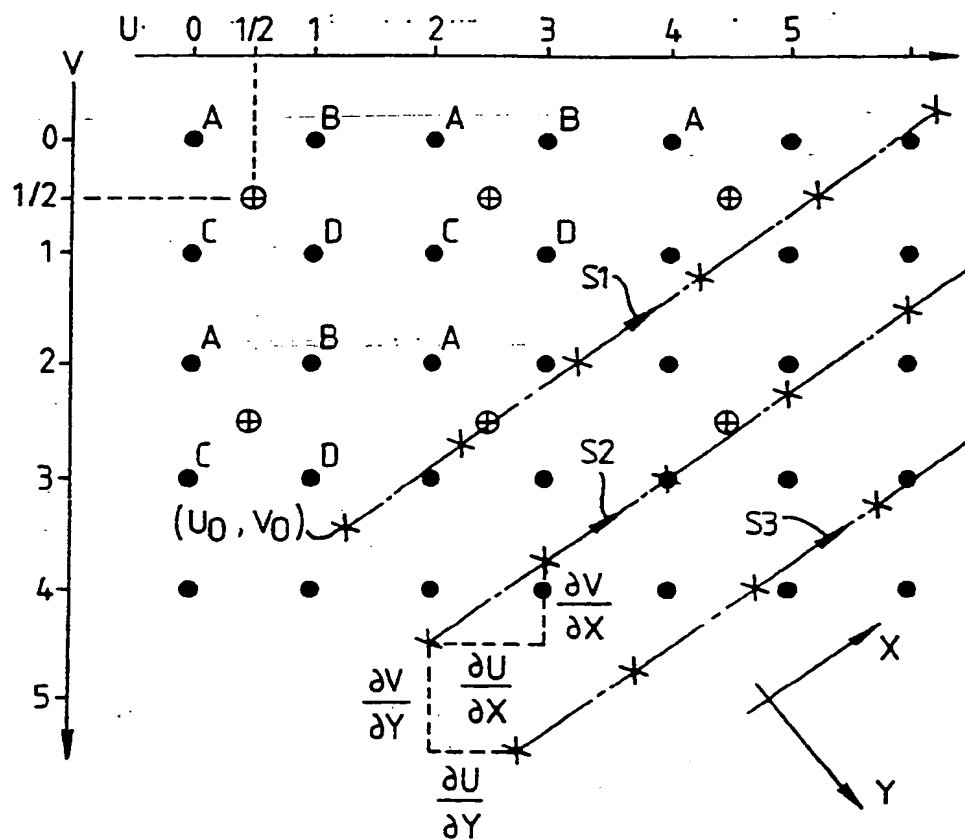


Fig. 7.

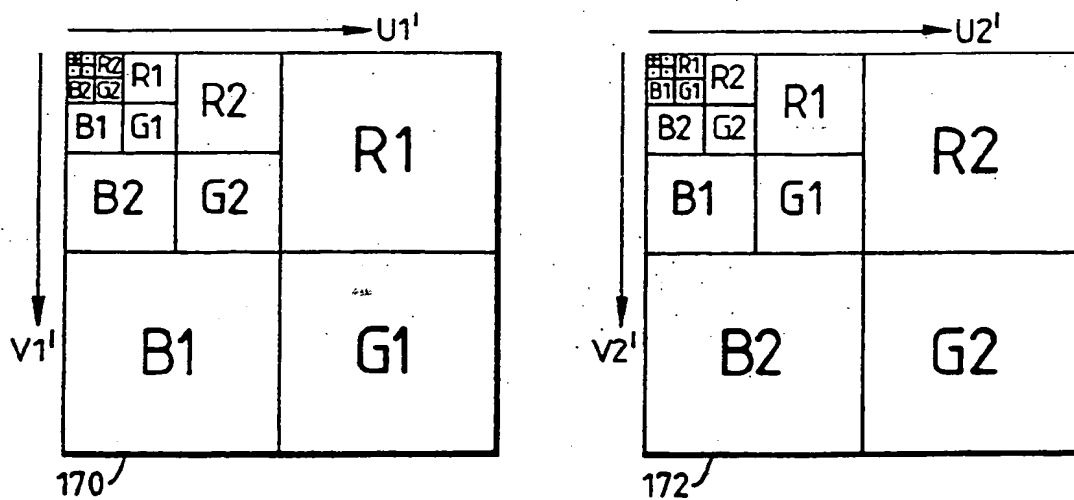
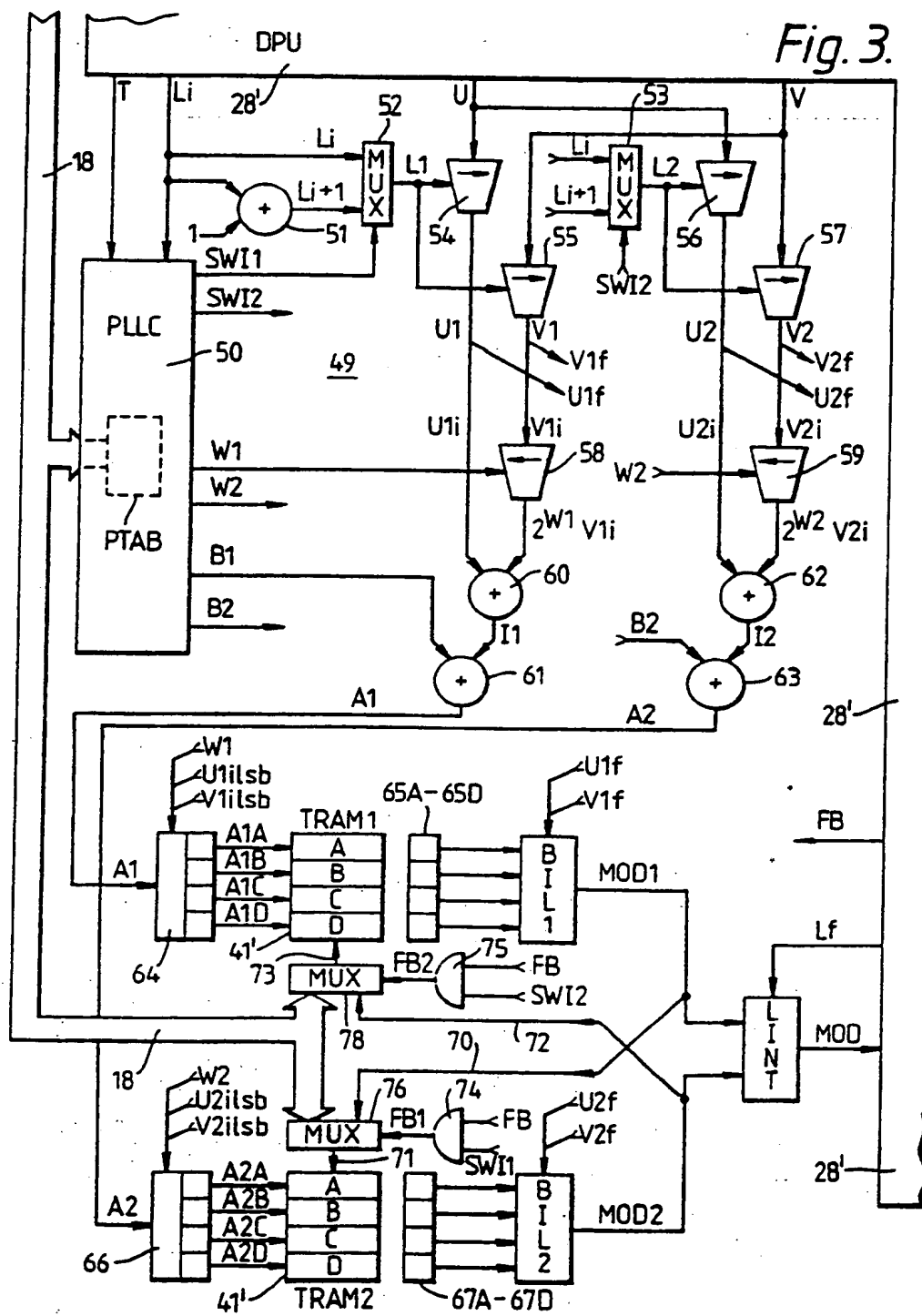
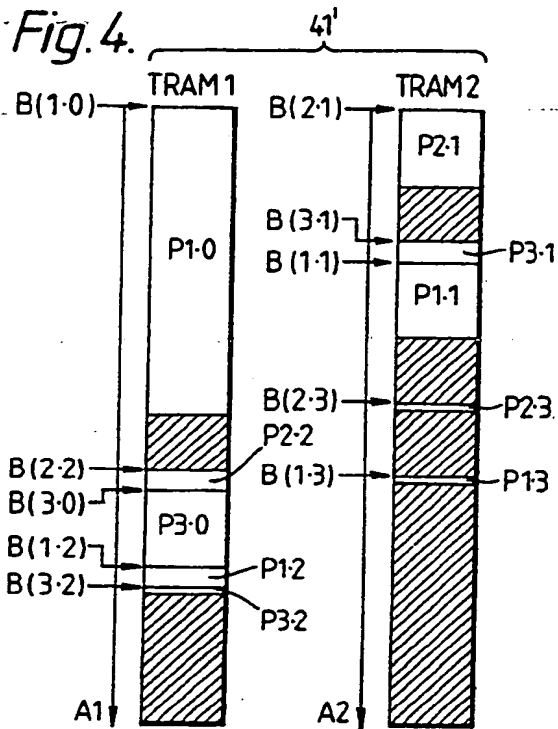


Fig. 3.

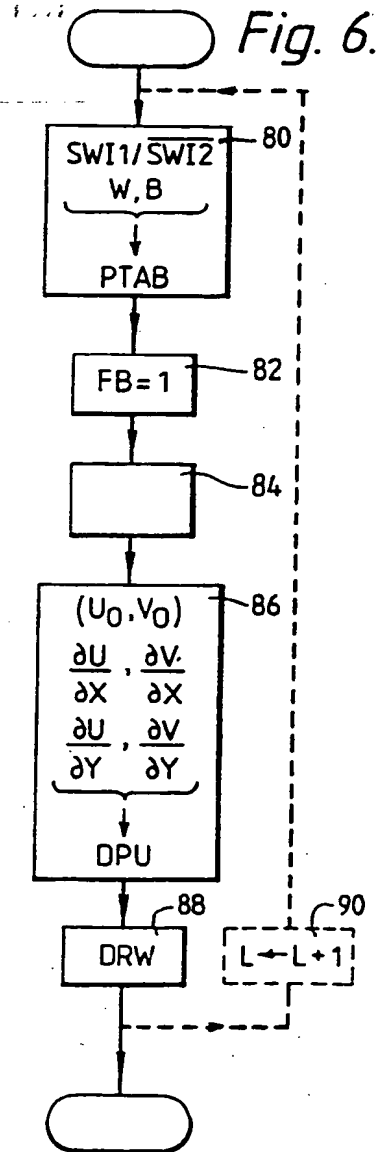




PTAB

T.Li	SWI1/SWI2	W	B
1·0	1	9	B (1·0)
1·1	0	8	B (1·1)
1·2	1	7	B (1·2)
1·3	0	6	B (1·3)
⋮			
2·0	—	—	—
2·1	0	8	B (2·1)
2·2	1	7	B (2·2)
2·3	0	6	B (2·3)
⋮			
3·0	1	8	B (3·0)
3·1	0	7	B (3·1)
3·2	1	6	B (3·2)
⋮			

Fig. 5.



THIS PAGE BLANK (USPTO)